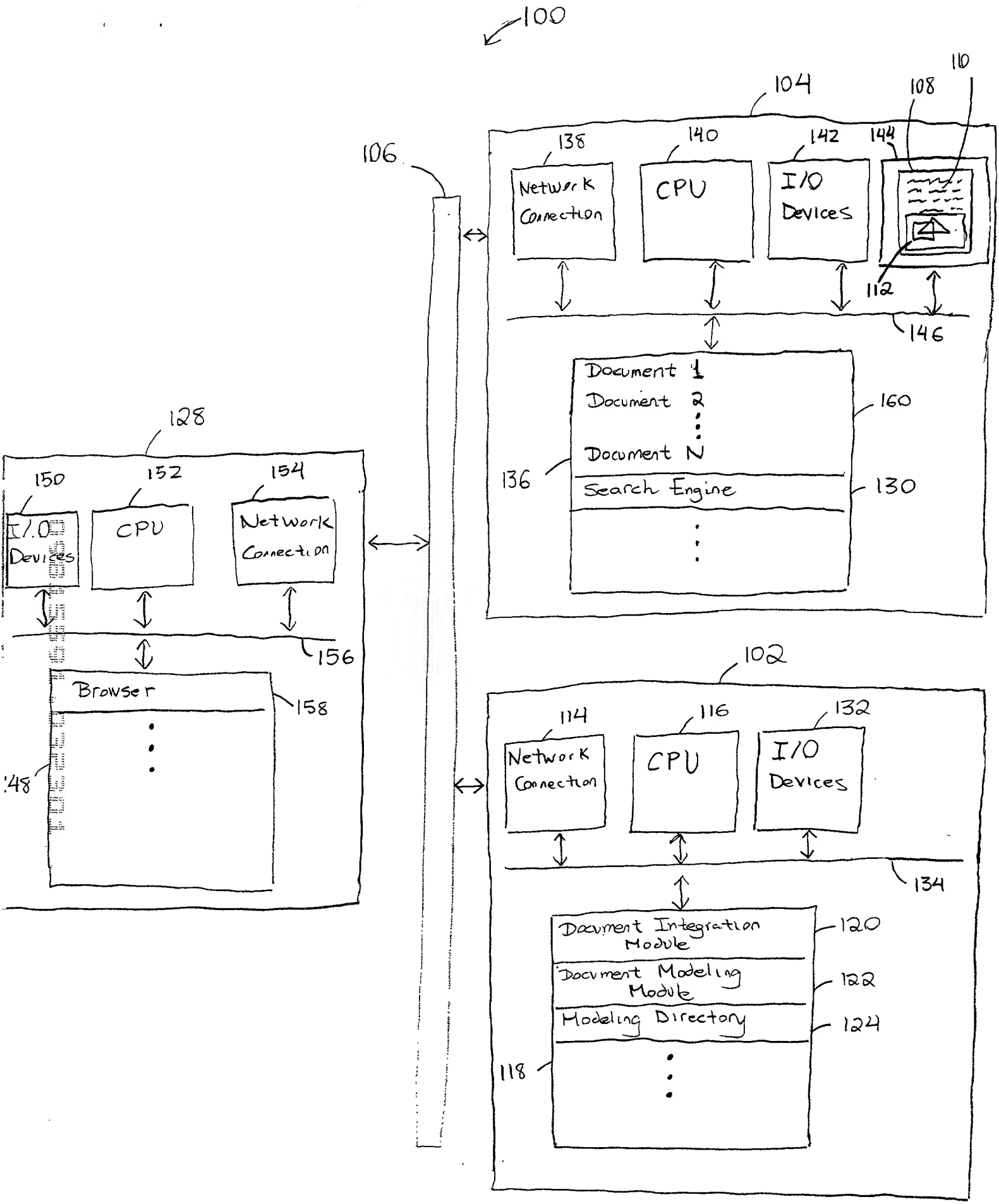

100
110
104
108
106
138
140
142
144
Network Connection
CPU
I/O Devices
112
146
Document 1
Document 2
Document N
160
136
Search Engine
130
128
150
152
154
I/O Devices
CPU
Network Connection
156
Browser
158
148
102
114
116
132
Network Connection
CPU
I/O Devices
134
Document Integration Module
120
Document Modeling Module
122
Modeling Directory
124
118

Fig. 1

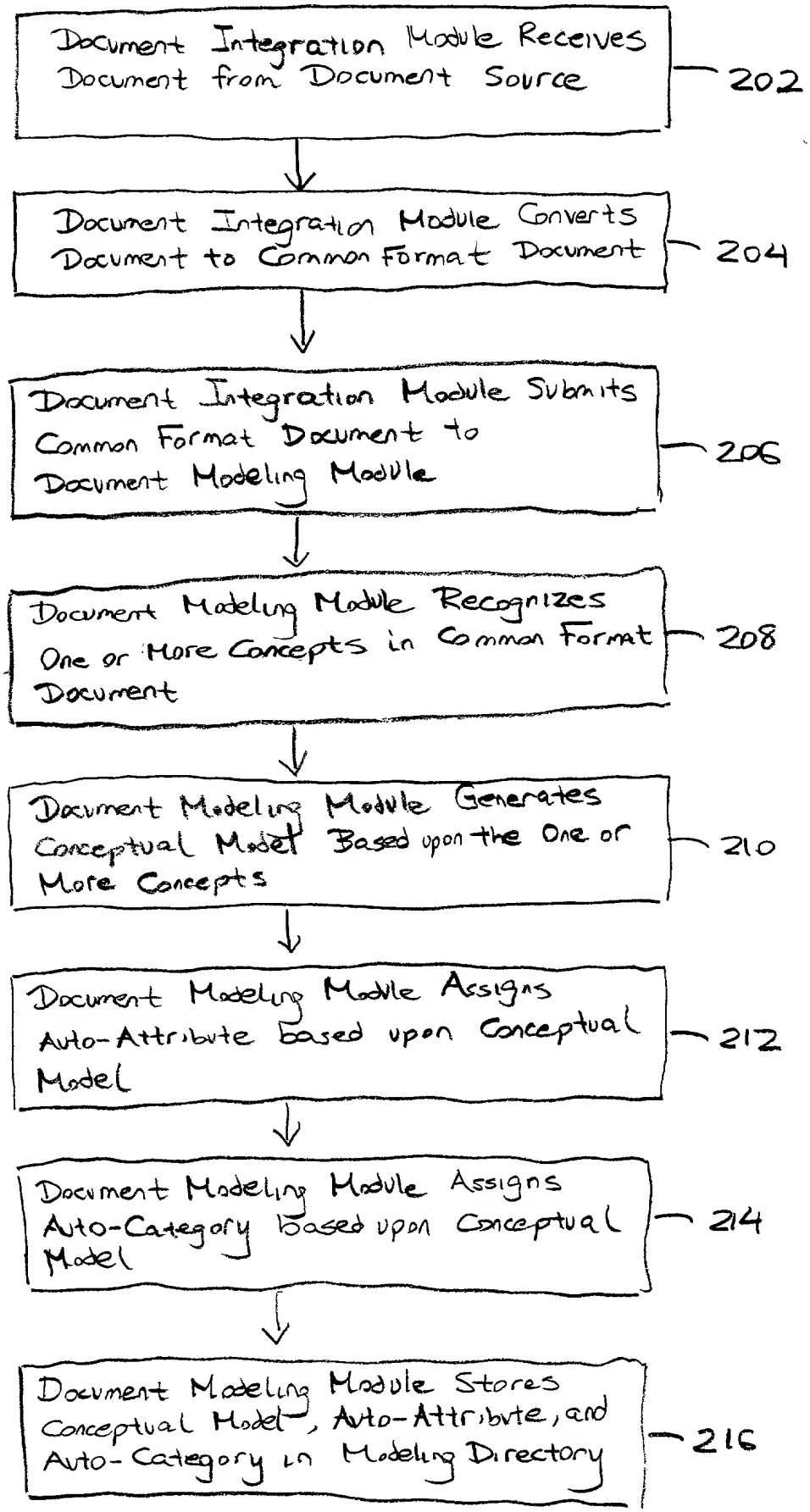

Document Integration Module Receives Document from Document Source
202
Document Integration Module Converts Document to Common Format Document
204
Document Integration Module Submits Common Format Document to Document Modeling Module
206
Document Modeling Module Recognizes One or More Concepts in Common Format Document
208
Document Modeling Module Generates Conceptual Model Based upon the One or More Concepts
210
Document Modeling Module Assigns Auto-Attribute based upon Conceptual Model
212
Document Modeling Module Assigns Auto-Category based upon Conceptual Model
214
Document Modeling Module Stores Conceptual Model, Auto-Attribute, and Auto-Category in Modeling Directory
216

Fig. 2

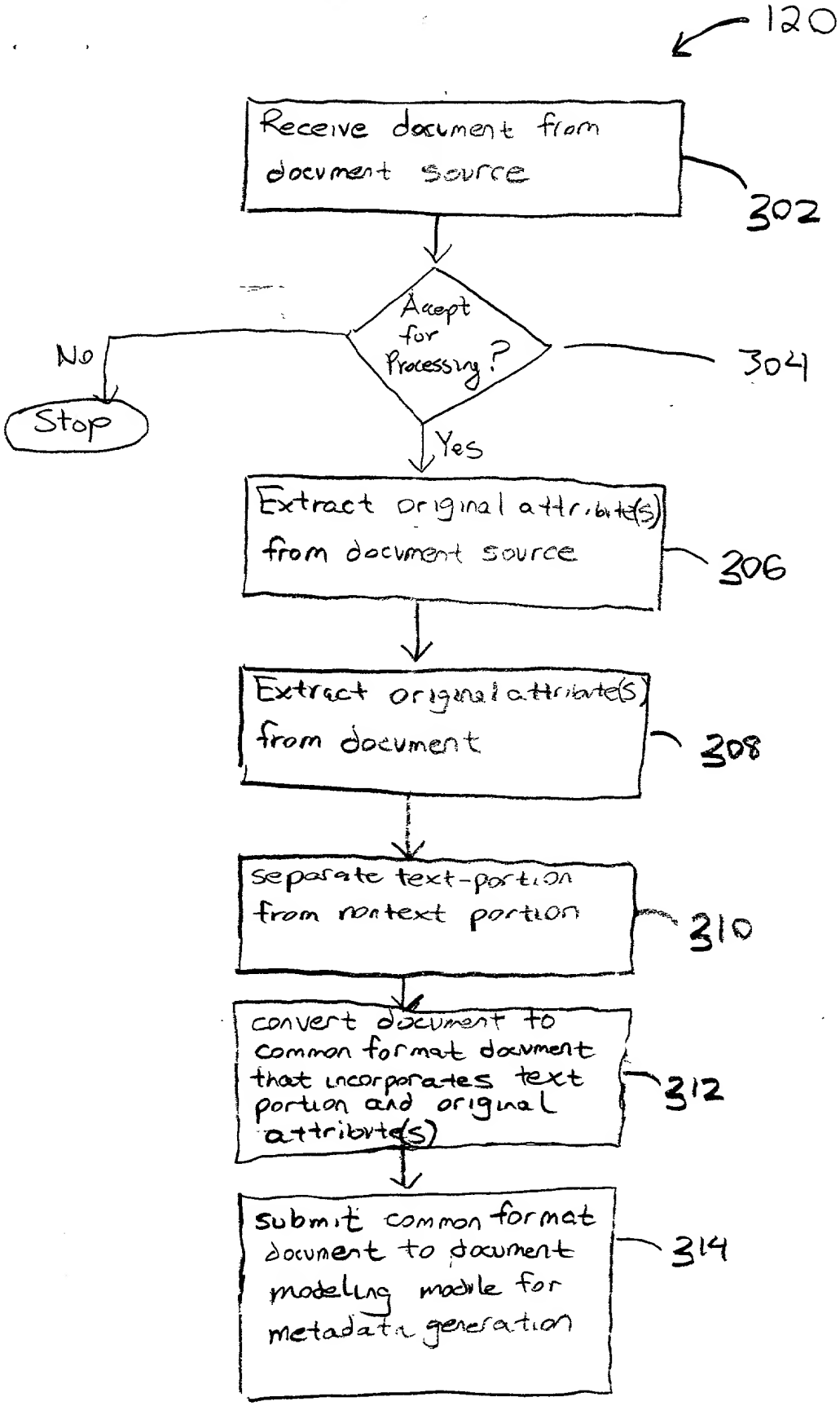

120
Receive document from document source
302
Accept for Processing?
304
No
Stop
Yes
Extract original attribute(s) from document source
306
Extract original attribute(s) from document
308
separate text-portion from nontext portion
310
convert document to common format document that incorporates text portion and original attribute(s)
312
submit common format document to document modeling module for metadata generation
314

Figure 3

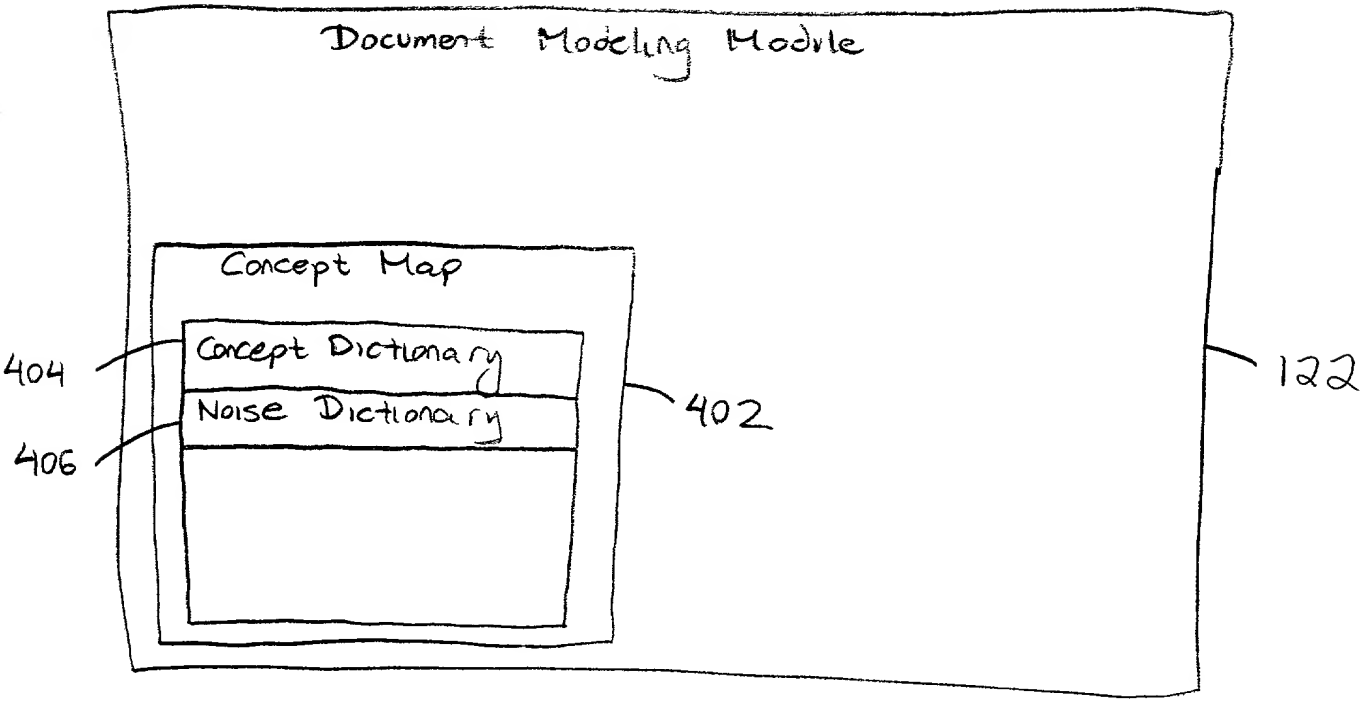

Document Modeling Module
Concept Map
Concept Dictionary
Noise Dictionary
404
406
402
122

Fig. 4

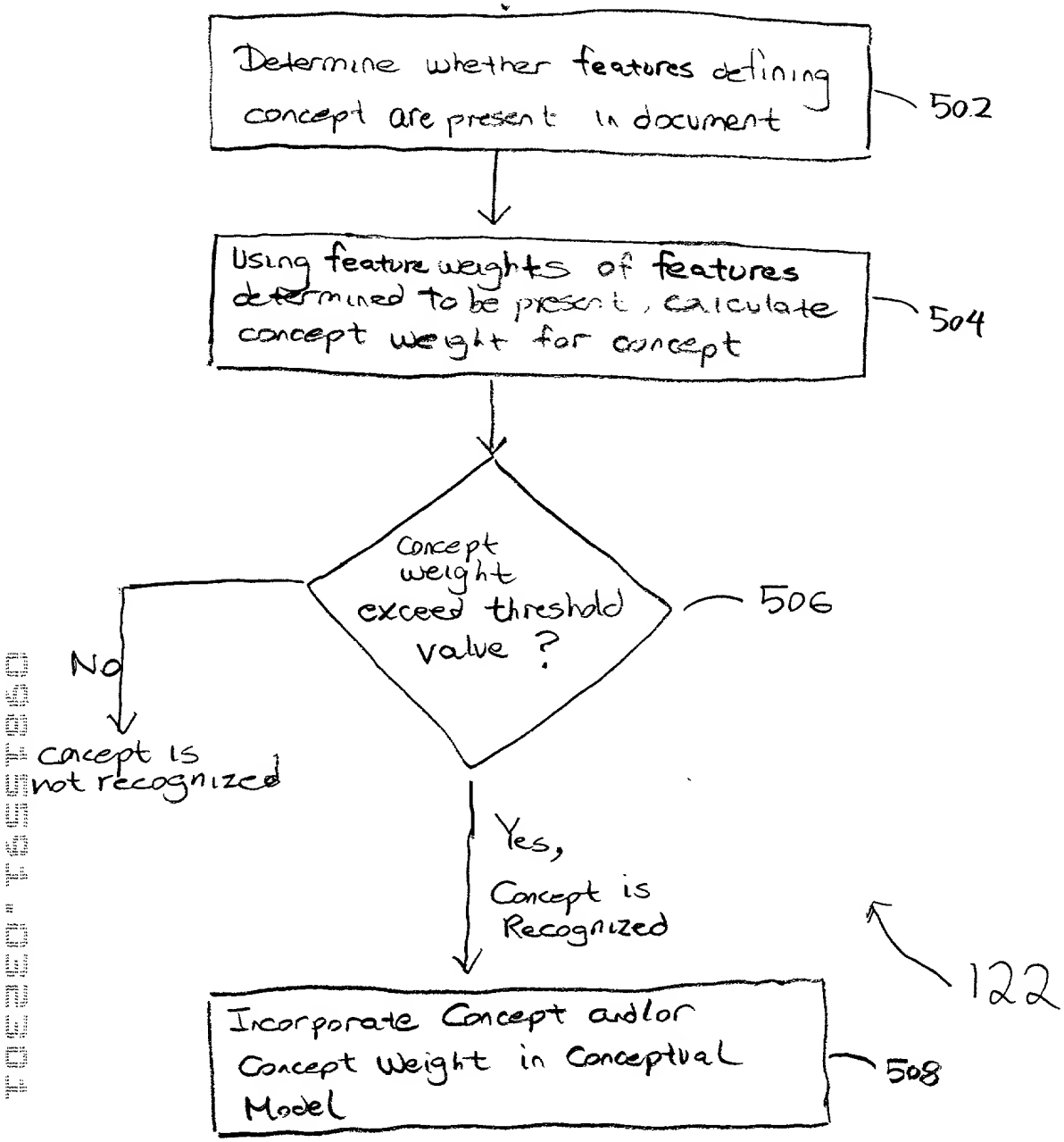

Determine whether features defining concept are present in document
502
Using feature weights of features determined to be present, calculate concept weight for concept
504
Concept weight exceed threshold value ?
506
No
Concept is not recognized
Yes,
Concept is Recognized
122
Incorporate Concept and/or Concept Weight in Conceptual Model
508

Fig. 5

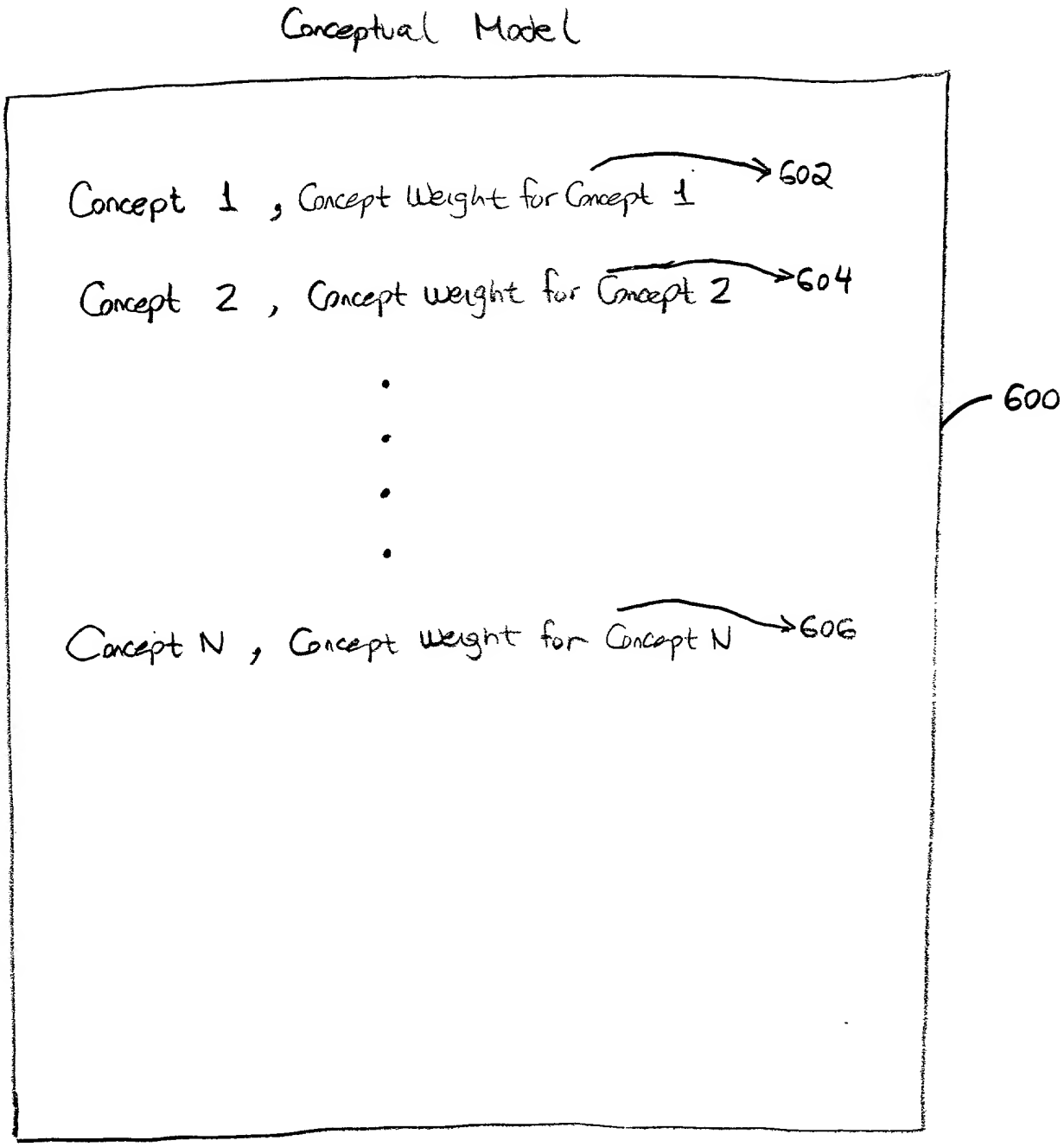

Conceptual Model
Concept 1 , Concept Weight for Concept 1
602
Concept 2 , Concept Weight for Concept 2
604
600
Concept N , Concept Weight for Concept N
606

Fig. 6

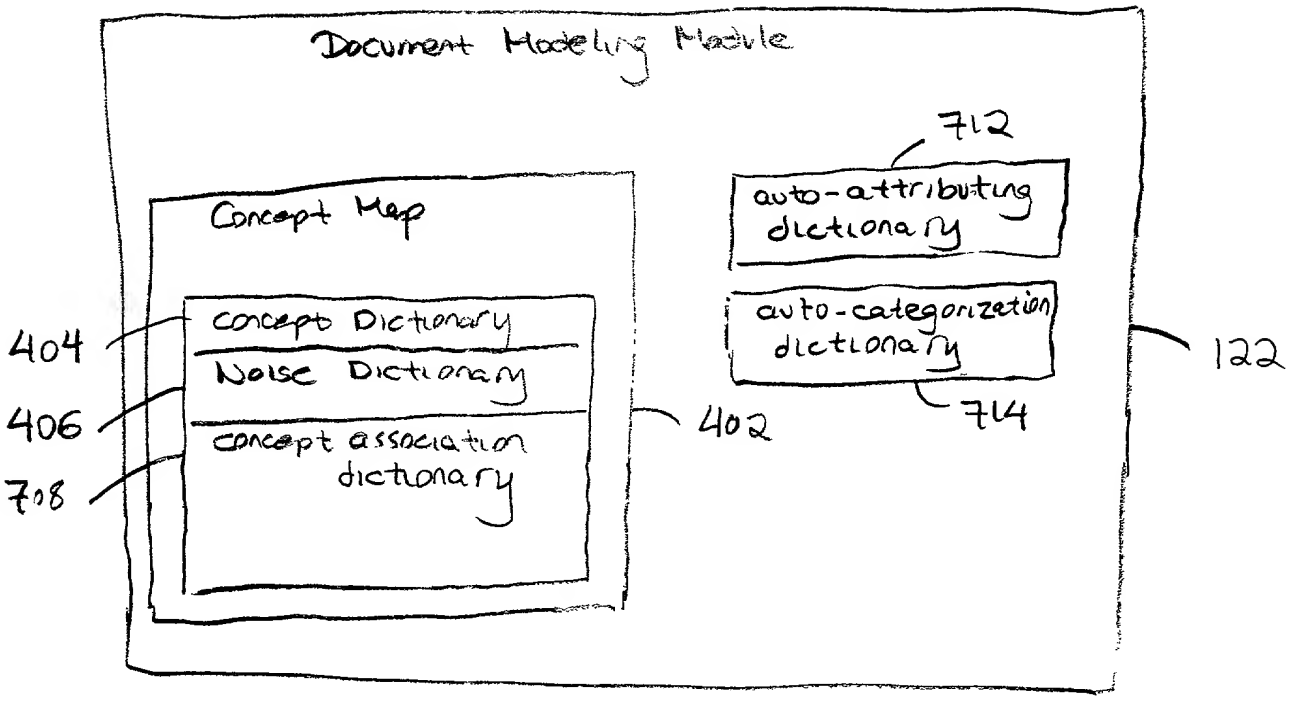

Document Modeling Module
712
auto-attributing dictionary
Concept Map
Concept Dictionary
Noise Dictionary
Concept association dictionary
auto-categorization dictionary
404
406
708
402
714
122

Fig. 7

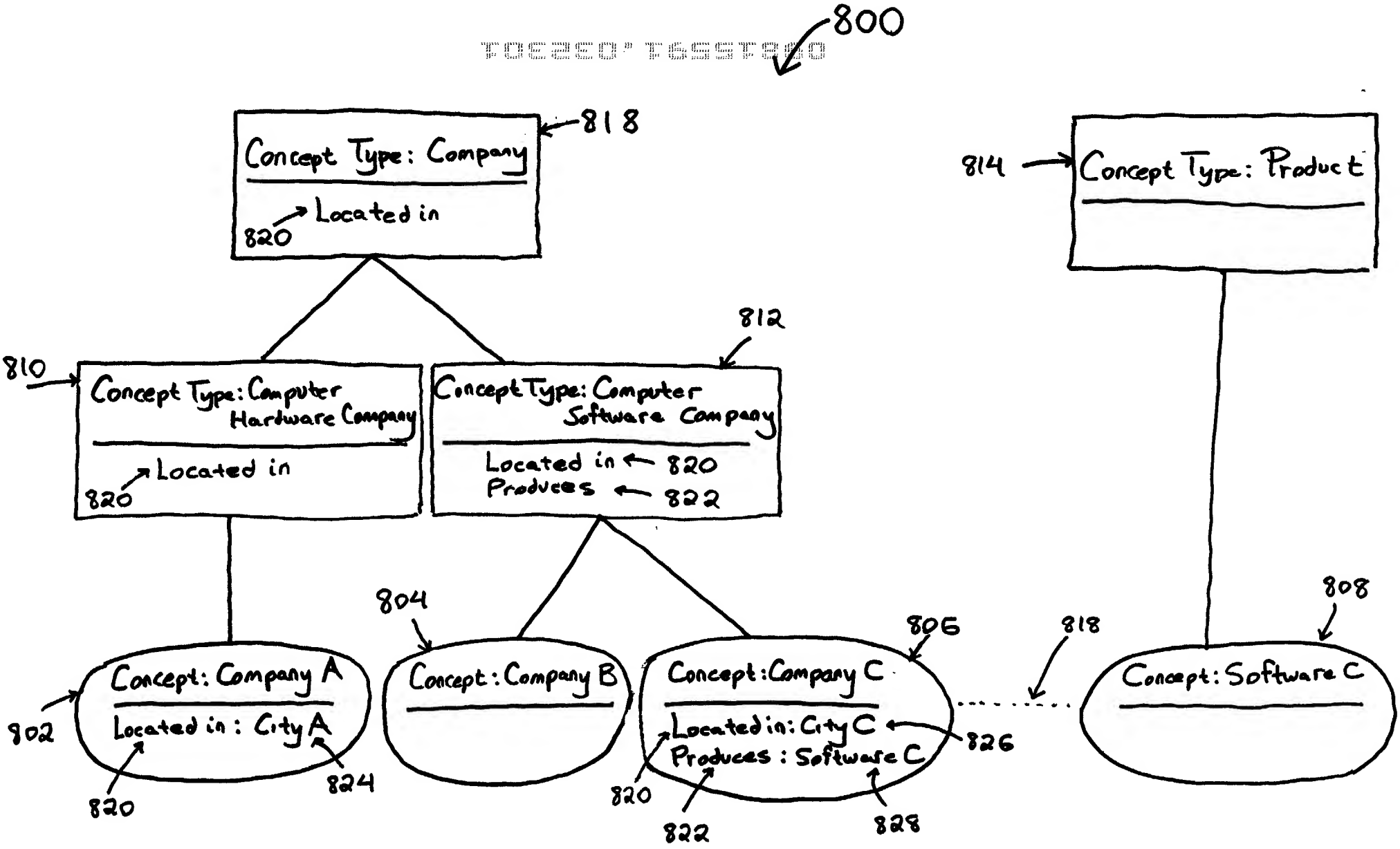
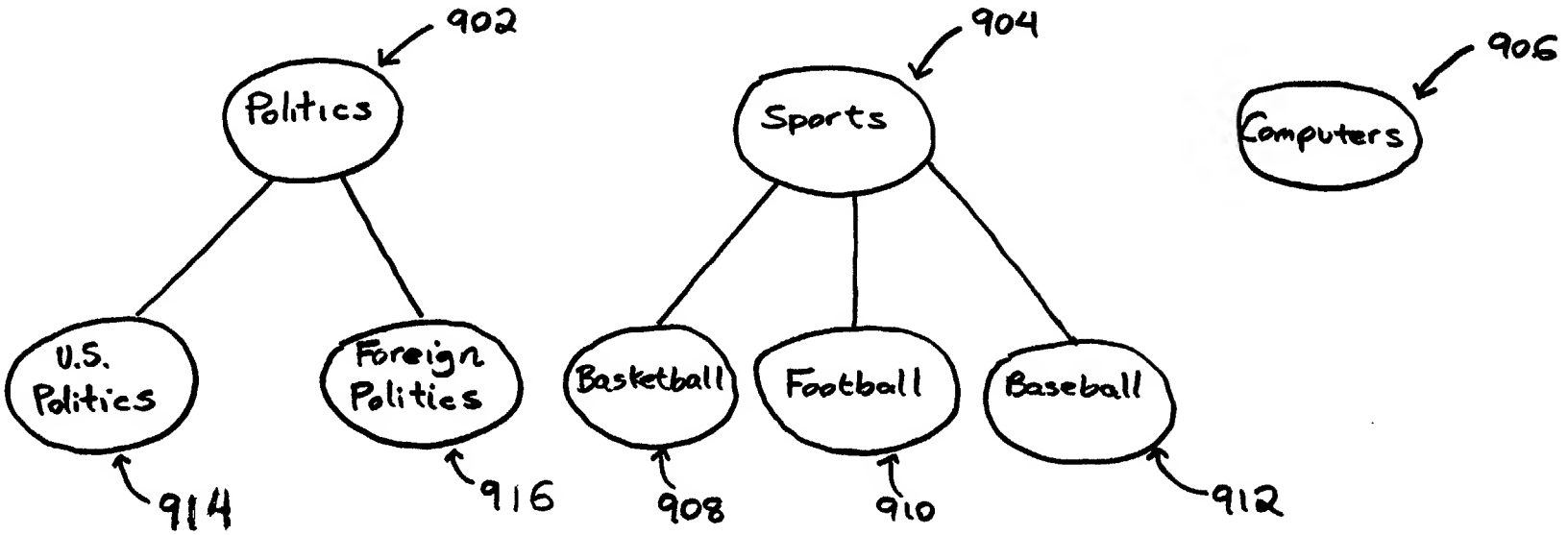

800
Concept Type: Company
818
820
Located in
814
Concept Type: Product
812
810
Concept Type: Computer Hardware Company
820
Located in
Concept Type: Computer Software Company
Located in ← 820
Produces ← 822
804
806
818
808
Concept: Company A
Located in : City A
Concept: Company B
Concept: Company C
Located in: City C
Produces : Software C
Concept: Software C
802
820
824
820
822
828
826

Fig. 8


902
Politics
904
Sports
906
Computers
U.S. Politics
Foreign Politics
Basketball
Football
Baseball
914
916
908
910
912

900

Fig. 9

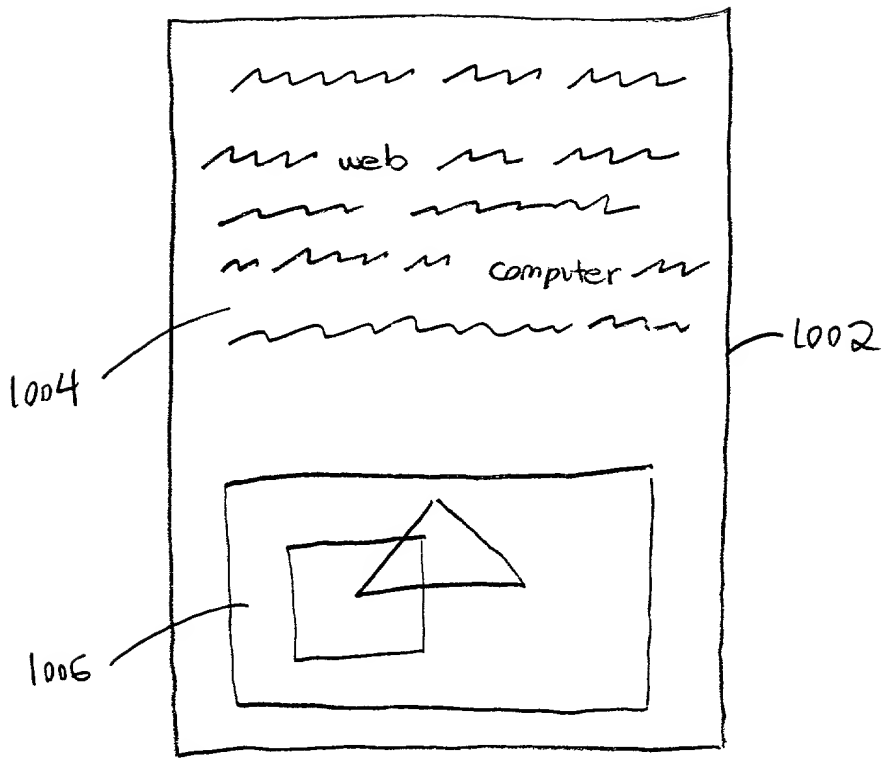


Fig. 10A

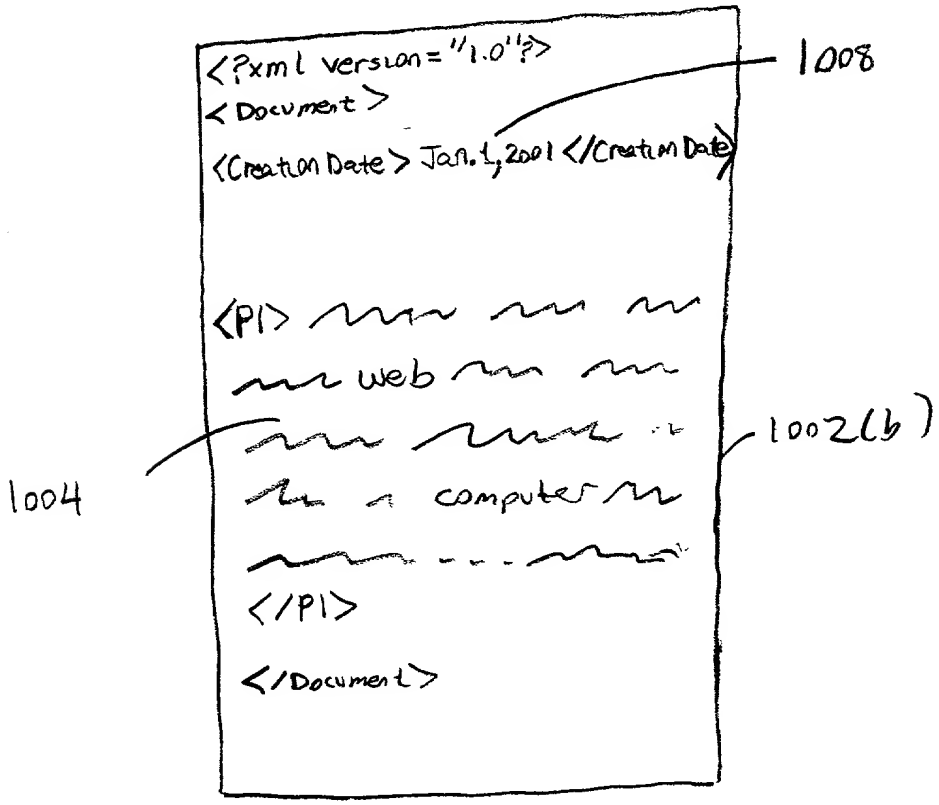


Fig. 10B

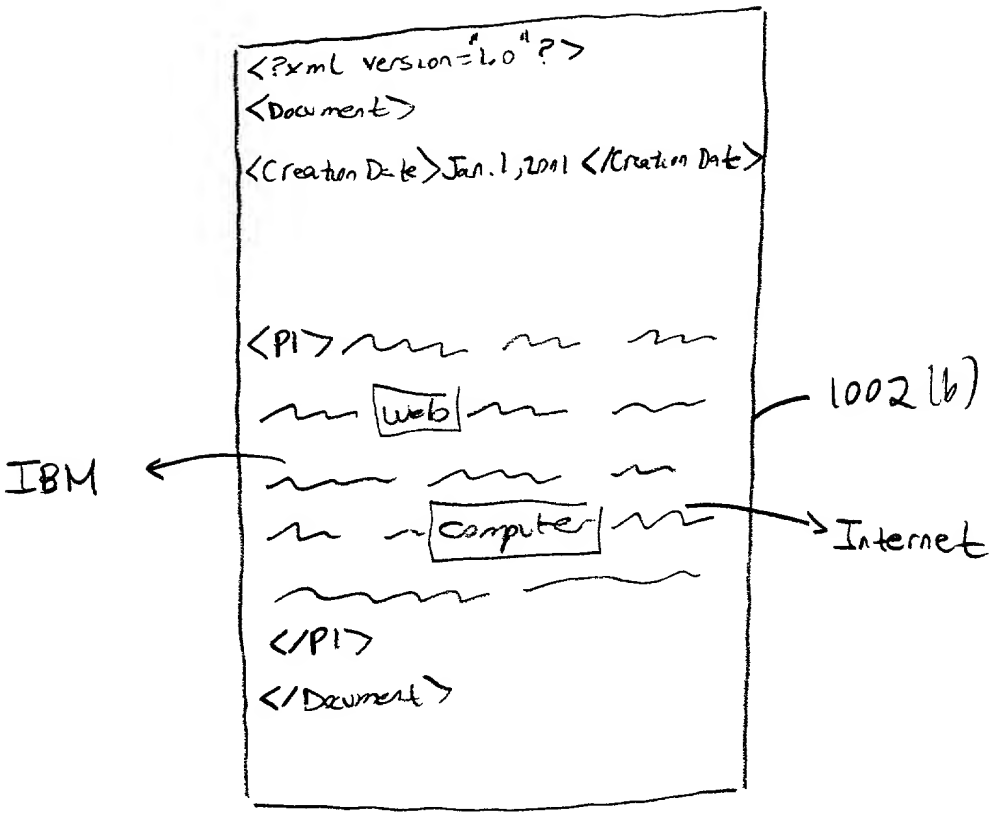


Fig. 10C

Internet	0.90
IBM	0.52

1010

Fig. 10D

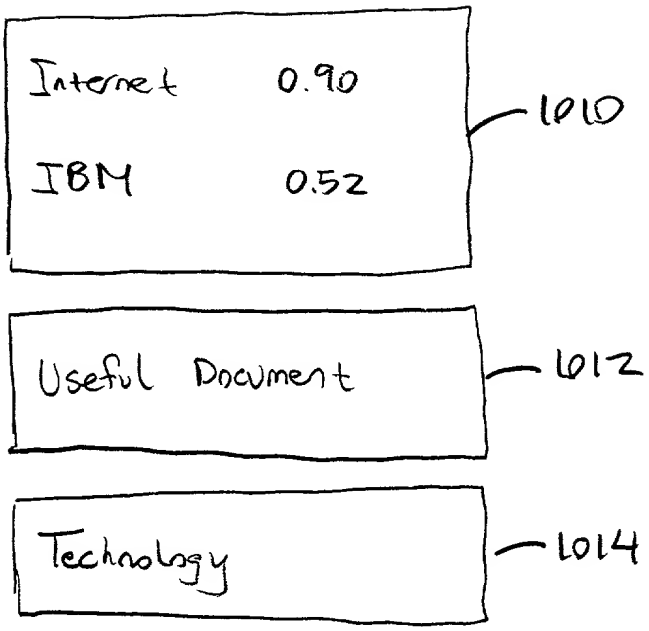

Internet 0.90
IBM 0.52
1010
Useful Document
1012
Technology
1014

Fig. 10E